DOCUMENT RESUME

ED 248 143 SE 044 875

AUTHOR Robertson, Craig A.
TITLE An Introduction to Chemical Abstracts, with Exercises.
PUB DATE [Aug 84]
NOTE 33p.; For other workbooks in this series, see SE 044 876-877.
PUB TYPE Guides - Classroom Use - Materials (For Learner) (051)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Abstracts; *Chemical Nomenclature; *Chemistry; *College Science; Higher Education; *Indexes; *Information Retrieval; Science Education
IDENTIFIERS *Chemical Abstracts

ABSTRACT

This workbook is the first in a series of three that has been integrated into the chemistry curriculum for majors at the University of Vermont. The workbook consists of exercises designed to provide undergraduate students with foundation skills in the use of professional literature and a familiarity with the printed "Chemical Abstracts." Each exercise consists of discussion material and instructions and a fill-in worksheet. Exercise topics focus on: (1) the nature of "Chemical Abstracts" (CA); (2) CA Author and Patent Indexes; (3) CA General Subjects Index; (4) CA Chemical Substances Index; (5) chemical substance names in the Index Guide; (6) chemical substance names and the Formula Index; and (7) chemical substance names and the Index of Ring Systems. A 6-item student evaluation sheet is included. (JN)

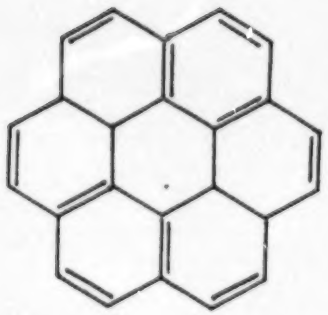
****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

NAME:_____ No:_____

ED248143

AN INTRODUCTION TO CHEMICAL ABSTRACTS

WITH EXERCISES



U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Craig A. Robertson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BY: CRAIG A. ROBERTSON

SE044875

Dear Reader:

This workbook is the first in a series of three that we have integrated into the chemistry curriculum for majors at the University of Vermont in order to train undergraduates in the use of professional literature. This workbook aims to provide them with foundation skills and a familiarity with the printed Chemical Abstracts. It is introduced in the first semester of a B.S.-track course in Organic Chemistry, after students have acquired some knowledge of IUPAC rules of chemical nomenclature.

The workbook's method emphasizes practical approaches to problems of the controlled vocabularies. The use of multiple fill-in sets for the questions encourages students to work independently of one another. One set has been filled in for the purpose of illustration.

Later workbooks in the series assume the completion of this one. They are:

An Introduction to the Literature of Organic Preparations
The Seminar Paper: Strategies for the Literature Search

For Angus and Sandy

Knowing names is my job. My art. To weave the magic of a thing, you see, one must find its name out . . . . There are spells to learn, of course, ways to use words; and one must know the consequences too. But what a wizard spends his life at is finding out the names of things, and finding out how to find the names of things.

Ursula LeGuin, The Tombs of Atuan

Directions for the Exercises

1. Write the number of your fill-in sheet in the blank at the top right corner of the workbook cover.

2. Fill in the blanks in the questions with the words, phrases or structures that complete the meaning of each.

3. You may need to visit another library on campus or learn about its holdings in order to answer some of the questions.

4. Work independently. If you get stuck on a question, you should re-read the text in the workbook that deals with that feature of Chemical Abstracts. If you suspect an error in a question -- always a possibility -- ask about it at the desk.

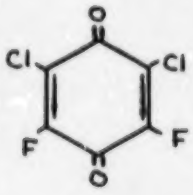
5. Please, do not make marks in the indexes or abstracts.

6. When you have finished the questions, turn in the workbook at the desk of the P/C Library. It will be corrected and returned at the table in front of the desk, usually within 48 hours. You must correct any errors and resubmit the workbook. Credit is given when all answers are correct.

Workbooks must be finished by: _____

No. _____

Fill-in Sheet

1. Biochem Z
2. May 7, 1984; minovincine
4. Seventh Collective Index; Allen, Christopher W.
6. Neutron activation analysis
7. Free energy; Free energy of activation
8. Anthocyanidins, anthoxanthins and betalins
9. Ecotrin
10. Remerin
10b. NMR data
12. 

13.
13a. a carboxylic acid

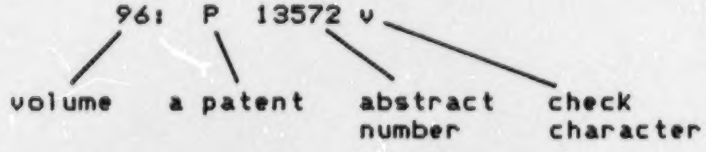
The discipline of chemistry shares with some other sciences -- most notably with the biological and medical sciences -- a preoccupation with names. Names of reactions, of chemical and physical phenomena, of publications, of procedures, and most especially, of chemical substances. As we begin to wade into the professional literature of chemistry in order to find information of value to our research or coursework interests, we will have to confront this bewildering vocabulary. Browsing through journals is an absurdly ineffecient way to find that information. To make best use of our time, we must use abstracts or indexes. But when each research article is indexed, it is, in a fashion of speaking, squeezed of its contents until only a small set of index names remain by which we may find it. An indexer chose these names from an official, "controlled vocabulary" or assigned them according to strict rules. In this workbook it is the controlled vocabulary and strict rules of nomenclature of Chemical Abstracts that interest us. We will not master them cheaply. But for almost every search for literature by means of Chemical Abstracts we must start by finding correct index names for the thing that we pursue.

Chemical Abstracts ("CA") celebrated its 75th anniversary of publication in 1982. It is the survivor of a number of national indexing systems -- English, German, etc. -- and has become the premier guide to the literature of chemistry, biochemistry, chemical engineering, pharmacology and other chemical sciences. The subject matter of CA is new science and engineering research. It does not report "popular" literature that rehashes what is common scientific knowledge or articles that deal with the purely commercial activitities of the chemical industries.

Every year CA abstracts nearly half a million journal articles, patents, dissertations, books, conference papers and other documents. These abstracts are descriptive summaries with bibliographic references that lead us to the primary research literature. They are published weekly and appear in 80 subject sections. Sections 1-34 alternate weeks with sections 35-80. Over the years there have been changes in the numbering and content of sections, but as of January 1982 (11th Collective Period) they were arranged thus:

1-20	Biochemistry; Education; Documentation
21-34	Organic Chemistry (including amino acids, peptides, carbohydrates)
35-46	Macromolecular and Polymer Chemistry
47-64	Applied Chemistry & Chemical Engineering
65-80	Physical & Analytical Chemistry

Each abstract starts with an abstract number, the first two digits of which (before the colon) indicate the volume of CA. The final character is a letter, which has no use to us in recent volumes of CA, but in earlier volumes it indicates the area of a column in which the abstract is located. An abstract may also contain a capital R, B, or P, which indicate a Review, a Book or a Patent respectively:



The title of the article or chapter appears in bold-faced type, followed by the author(s). The title of the journal or proceedings appears in Italic type and is almost always abbreviated. The key to these abbreviations is the Chemical Abstracts Service Source Index ("CASSI"). CASSI is arranged alphabetically -- by the

abbreviations, which can be confusing. The abbreviations have become standard in citing chemical publications. CASSI provides full titles and other publishing data. Because it lists about 100,000 active or discontinued titles, many abbreviations look similar, so it is good practice to look up the abbreviations. Don't guess. CASSI consists of 2 cumulated volumes for the years 1907-1979, with quarterly supplements for new or changed titles. Begin by looking in the main set and move on to the supplements if necessary.

Each weekly issue of Chemical Abstracts contains a kind of provisional "Keyword Index" to that issue. This index uses words taken from the authors' titles to guide us to the abstracts. These keywords are not consistent and do not come from the controlled vocabulary. So they are of value for the individual issues only, and are superseded by the volume indexes when these appear.

Each "volume" of Chemical Abstracts consists of 26 weekly issues. When a volume is complete, Chemical Abstracts Service provides a set of volume indexes for it:

Index	Color Key @ UVM
Author Index	White
General Subjects Index	Blue
Chemical Substances Index	Green
Patent Index	Black
Formula Index	Yellow
Index of Ring Systems	
Index Guide	Red

At the end of every 5-year period (10 years for the early volumes) these volume indexes are cumulated into a set of Collective Indexes. The Collective Index Periods of Chemical Abstracts are:

11th Collective Index	Vols. 96-105	1982-1986
10th Collective Index	Vols. 86-95	1977-1981
9th Collective Index	Vols. 76-85	1972-1976
8th Collective Index	Vols. 66-75	1967-1971
7th Collective Index	Vols. 56-65	1962-1966
6th Collective Index	Vols. 51-55	1957-1961
5th Decennial Index	Vols. 41-50	1947-1956
4th Decennial Index	Vols. 31-40	1937-1946
3rd Decennial Index	Vols. 21-30	1927-1936
2nd Decennial Index	Vols. 11-20	1917-1926
1st Decennial Index	Vols. 1-10	1907-1916

1. What is the full title of a publication with the CA abbreviation: Biochem Z. ?

ANSWER: _____

_____

1a. Does UVM own it?

ANSWER: _____

1b. If yes, where is it located?

ANSWER: _____

2. Consult the CA issue for the week of May 7, 1984

Find the title of the first article on minovincine

ANSWER: _____

2a. Does UVM receive (or own) the publication in which it appears?

ANSWER: _____

3. Is Chemical Abstracts an appropriate index to find articles on the costs of petrochemical feedstocks to the chemical industry?

ANSWER: _____

3a. Explain your answer to 3.

ANSWER: _____

_____

_____

The Author and Patent Indexes

The Author Index of Chemical Abstracts resembles other author indexes and holds no special mysteries. Remember that either individuals or organizations (e.g., Exxon Research and Engineering) may appear as authors. The only name problem that we may meet in this index results from the fact that in the age of Big Science research groups rather than heroic individuals, like Dalton or Faraday, conduct most research, so most publications have multiple authors. The Author Index provides an abstract number reference only for the first author mentioned in the paper. All other authors are listed with cross-references to the first author:

Watson, J.D. See Crick, F.H.C.

A descriptive ("text modification") phrase, which follows closely the words of the title, appears with each abstract reference.

The Patent Index is an index and concordance only to patent numbers. It cannot be searched by subject, so we will not describe it in detail. It provides abstract references for each original patent. It also serves as a concordance, linking foreign equivalent patents to the original patent and its CA abstract. Before 1981 the Numerical Patent Index (1937-1980) and the Patent Concordance (1962-1976) appear as separate indexes.

4. Find the first article referenced in 7th Collective Index that was authored by Allen, Christopher W.. Give a full bibliographic reference (all authors, title, full journal or book title, volume, pages and date) for the paper.

ANSWER: _____

_____

_____

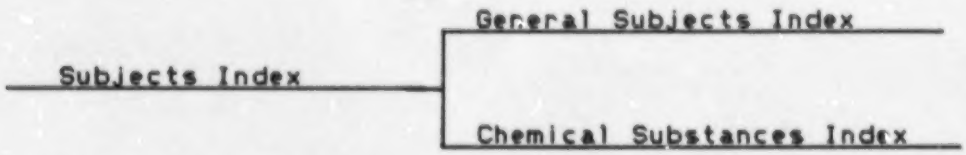
_____

4a. Does UVM own this publication?

ANSWER: _____

The General Subjects Index

The General Subjects Index to CA has appeared as a separate publication since 1972 (beginning of the 9th Collective Index Period). Before 1972 general subjects were interfiled with chemical substance names in a Subjects Index:



Chemical Abstracts treats as "general subjects" just about every index topic that it cannot classify as a specific chemical substance with a known composition. So, the General Subjects Index is the place to hunt for references on:

Concepts:	Entropy, Pi-bond
Classes of compounds:	Alkaloids, Alkyl bromides
Biological organisms and materials:	Aloe vera, Interferons
Reactions:	Alkylation, Diels-Alder
Lab & instrumental methods:	Chromatography, gas
Industrial products & processes:	Acrylic fibers, Cracking
Physical & chemical phenomena:	Density, Luminescence

It is obvious that much that is of particular interest to analytical, physical and industrial chemistry is best accessed through the General Subjects Index.

Indexers assign the general subjects names from a controlled vocabulary. The thesaurus for this vocabulary of names is the Index Guide. We should begin every search of the General Subjects Index by first examining the Index Guide. There is a separate Index Guide for each collective index period from the 8th. Each lists subject headings in bold-faced type. The IG may refer us

from a term that we had in mind to a correct index term by use of a mandatory ("See . . .") cross-reference:

> Phosphazenes:
> See Phosphonitrile Compounds

Or it may provide us with an explanatory scope note, or cross-references to related ("See also . . .") subjects, or both:

> Semiconductors, electric:
> Studies on materials used as electric semiconductors are indexed at this heading. See also headings for common semiconductor materials, such as Germanium, Indium antimonide . . . .

To trip up the unwarry user, the General Subjects Index has a number of traps. It is best for us to know about these pitfalls from the start. First, although the Index Guide is an essential first step, it does not list every subject heading. Chemical Abstracts Service seems simply to assume that some subject names are too obvious to include, while others -- especiallly biological names -- are so numerous that they would, if included, make the IG overly long. So the Index Guide is an essential starting place, but we should not draw a firm conclusion from its silence. A second important point about the General Subjects indexing vocabulary that is not obvious at first is its hierarchical structure, which is concealed by the alphabetical arrangement of the index. Indexing terms have a structured relationship to one another of more general and more specific subjects. The importance of this hierarchy lies in the Chemical Abstracts policy that requires the indexers to use the most specific terms available to

index each article. We need to be aware of this practice and to plan our search of the General Subjects Index so that we start with the most specific term used that will cover the meaning of our subject. If it, or a suitable synonym is not used in the index, we may choose a more general subject name from the "Hierarchies of General Subjects Headings". For example, we would look for articles on reaction kinetics

Affinity
. Reactivity
. . Electrophilicity
. . Nucleophilicity
. . Reducing power
. . Reaction mechanism
. . . Kinetics, reaction

under "Kinetics, reaction" (more specific) rather than, say, "Reactivity" (more general). The Index Guide (Appendix I) lists these "Hierarchies of General Subjects Headings" and even provides an index to them.

A third policy that we need to be on guard about in the IG and the General Subjects Index is the use of "class" headings to cover three or more different members of any group. For example, if we are interested in all binary halogen salts of Beryllium, we need to use

Beryllium halides

in the General Subjects Index, as well as Beryllium Chloride, Beryllium Bromide, etc. in the Chemical Substances Index. Such "class" headings also appear in the "Hierarchies of General Subjects Headings" in Appendix I of the Index Guide.

5. In the last example, why are the individual halides found in the Chemical Substances Index?

ANSWER: _____

The following questions should be answered from the 10th Collective Period Indexes and Index Guide

6. If you are interested in finding literature on neutron activation analysis what general subjects should you use for the 10th Collective Index Period?

ANSWER: _____

6a. Find the Chemical Abstracts volume and abstract number for the first review article on that subject in the 10th Collective Index.

ANSWER: _____

7. Would the index term free energy be an appropriate heading to choose for literature on free energy of activation?

ANSWER: _____

7a. Explain your answer.

ANSWER: _____

_____

8. If you wished to find information on anthocyanidins, anthoxanthins and betalins what "class" heading(s) might you use? (You may need to look at the "Hierarchies of General Subjects Headings" in Appendix I of the Index Guide).

ANSWER: _____

The Chemical Substances Index

In the Nineteenth Century and the earlier part of the this century the syntheses of new compounds and the discovery of new preparative reactions dominated the development of chemistry. New synthetic methods, isolation of natural products and laboratory development of new chemical compounds continue to be important concerns of the discipline. Chemical Abstracts indexing system has inherited an emphasis on specific chemical compounds, along with some problems that this approach presents. Nowhere in the use of CA is our struggle with names more demanding than in the recognition of individual chemical substances.

Any effective system of naming chemical substanes requires that:

1. the names be systemmatically derived from structures
2. each name be unique and unambiguous

To appreciate why this is so, consider the fact that almost 7 million chemical substances are known to the literature, and about 300,000 new ones enter it each year! To complicate matters, trivial and trade names abound in chemical communication. We have already learned something about the use of IUPAC rules to name compounds. We are probably able to apply the simpler rules to a few thousand substances. But the rules that we need to name 7 million substances unambiguously from their structures almost defy human mastery. Indeed, Chemical Abstracts Service uses a computer-based Registry System to verify the assignment of chemical names. The inevitable result is a system of rather

formidable "official" chemical names. These names are based generally on IUPAC rules, but they have actually grown stricter in application in recent years -- especially, at the outset of the 9th Collective Index Period (1972).

Chemical substance names have some analogies with human names in the way that they are listed in directories. For example, personal names appear in the phone book or campus directory last-name-first. Chemical compounds are also listed in an inverted way in CA, with the name of a parent compound appearing first. The parent compound takes its name from the basic chain or ring system and the principle functional group of the molecule:

Hexanedioic acid

Where we must consider two or more functional groups, a hierarchical list governs priority in naming the parent compound. The hierarchical list appears in Section 106 of the Index Guide. Generally, radicals, ions and acids rank high in this order of priority. The name of the parent appears first. Functional group derivatives follow. So, for example, we look for the dimethyl ester of Hexanedioic acid after the parent:

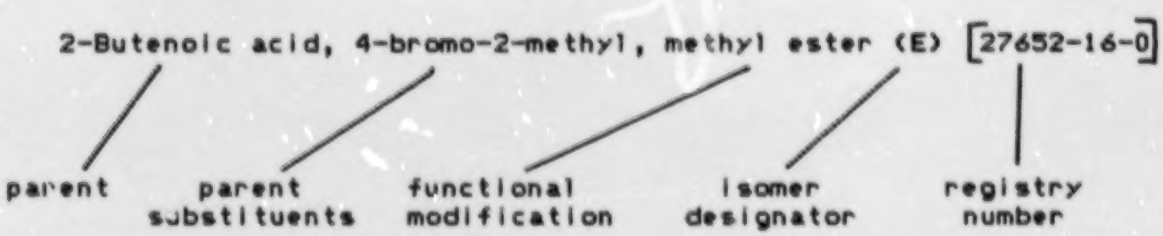
Hexanedioic acid [124-04-9]
_____, dimethyl ester [627-93-0]

If the parent compound is substituted, it in turn becomes the parent of a cadet (junior) branch of the family, following the unsubstituted parent and its derivatives in the index:

Hexanedioic acid, -2,5-dibromo [3479-85-4]

_____, dimethyl ester [868-72-4]

Isomer designators are an essential part of the systematic name.



The number that accompanies a name, called a Registry Number, should be considered part of it. These numbers convey no structural or chemical information, but have a different kind of analogy with human names. Like the "taxpayer identification numbers" or social security numbers that follow us through life, they are a shorthand way of identifying individual compounds and (almost) never change after they are first assigned by Chemical Abstracts Service. I may become a famous novelist and adopt a nom de plume, or I may change my name legally, but I will always be known to Bureaucracy by the same 9 digits. Likewise, chemicals can and do change names. Asprin is known by more than a hundred trade, trivial and quasi-systematic names -- but to CA it will always be 50-78-2. So the registry numbers help us follow a single substance through vagaries of name changes, and they are very useful for computer searching of chemical literature.

Look up a substance name in the Chemical Substances Index alphabetically, letter by letter, ignoring all breaks in the string of characters and all number or letter locants. Only when

two substance names differ only by the locants do we need to consider them, e.g.:

2,4-Hexadiene, 2,3-dimethyl before 2,4-Hexadiene, 2,4-dimethyl

One or more abstract references follow each substance name. Descriptive, "text modification", phrases, which indicate the content of each original document, accompany most abstract references. Abstract number references that apppear first, without such descriptive phrases, are usually for preparations:

8-Azabicyclo 3.2.1 octane-2-carboxylic acid, methyl ester 1R-(exo,exo) [50-36-2]
83: R152216b, R197782z; 85: P138339d
prepa- abuse of, clinical aspects of, 76: R108770c
rations acetic and acetoacetic acid incorporation, 77: 58902b
.
.
.

A specialist will master the rules that he or she needs to name a class of compounds that are of particular interest, but most of us must learn strategies -- tricks -- to discover the "true names" that will let us conjur information from the Chemical Substances Index. Three such strategies are:

1. Use of the Index Guide for trivially-named substances
2. Use of the (molecular) Formula Index, together with our passive ability to recognize correct and incorrect names when we see them
3. Use of the Index of Ring Systems as an entry into the Chemical Substances Index

The remainder of this workbook explains these three essential approaches to finding chemical substance information in Chemical Abstracts.

Chemical Substance Names in the Index Guide

Browsing through the Index Guide we notice that most of the entries (about 80% of them) give systematic names for substances. Many of these refer us from a trivial (i.e., any non-systematic) name to the correct CA index name. As numerous as these names in the IG may seem to us they comprise only a small fraction (100,000 or about 1.5%) of the vast number of compounds known to the Registry System. Happily, that small list includes many substances that are of intense interest, such as drugs, listed by their generic names, natural products, enzymes, polymers and common reagents. We may enter the Index Guide with a name that we find in some other source -- for example, in the Merck Index -- and locate the CA systematic name for it as it appears in the Chemical Substances Index.

For the following exercises use the Indexes of the 9th Collective Index Period.

(Hint: For generic drug names, use the Merck Index).

9. What is the generic name for ecotrin?

ANSWER: _____

9a. What is its systematic name in the 9th Collective Index?

ANSWER: _____

10. remerin is the trivial name of a chemical substance. What is its systematic name?

ANSWER: _____

10a. Give a full bibliographic reference (authors, title, full journal title, etc.) for a paper on its synthesis, preparation or isolation.

ANSWER: _____

_____

_____

10b. Find the article. Does it provide NMR data?

ANSWER: _____.

Chemical Substance Names and the Formula Index

A second friendly way of finding chemical substance names and literature references is to use the Formula Index. To start, we may find the molecular formula of a compound in a handbook, or we may derive it from our own knowledge of the structure. If we write out the formula from the structure, we should take special care in counting hydrogen atoms. Molecular formulas in the Formula Index and most handbooks appear in the Hill System, whose rules we may summarize briefly:

1. Elements appear in alphabetical order by symbol, except that . . .
2. If a molecule contains carbon, we count and write the carbon atoms first, and . . .
3. If it also contains hydrogen, we count and write the hydrogen atoms next.
4. Each element is listed in ascending numerical order (e.g., H_2O before H_3N)

The Hill System leads to some odd looking formulas, like ClNa or O_2S. But a more serious problem arises from the fact that a single molecular formula may represent many different compounds. The simple alkane formula, $C_{10}H_{22}$, has 75 possible structural isomers. But in most cases we will encounter a more manageable number of possibilities, so that by using our passive ability to recognize wrong names when we see them, and our partial knowledge of IUPAC rules, we can eliminate most of the candidates presented by the Formula Index, and pick the right name for the substance that interests us.

With each formula appear the names of substances, their registry numbers, and abstract references for each. For some

heavily cited substances we simply see a cross-reference to the Chemical Substances Index:

N_3P_3 1,3,5,2,4,6-Triazophosphorine [291-37-2] See Chemical Substances Index.

In any case, we should always take the compound name to the Chemical Substances Index, because it provides more information, in the form of the descriptive "text modification" phrases, which the Formula Index lacks.

Formula Index ----------> Chemical Substances Index

The arrangement of Hill formulas in an index or handbook considers each element symbol in alphabetical order, together with its numerical subscript in numerical order. This arrangement may be more felicitously shown than explained:

$AgNO_3$

CSe_2

$C_2H_4Br_2$

C_2H_5Br

HNO_3

H_2O

H_2O_2

H_3N

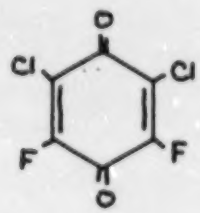
N_2O

11. Write the Hill System formula for each of the following, and arrange them in the order in which they would appear in the Formula Index: 2-bromobutanoic acid; 2-butanone; 1-chloro-2-butene; silver cyanide; 1,2-dibromobutane; potassium bromide.

ANSWER: _____

_____

12. The following is the structural formula of a simple compound. Write its molecular formula in the Hill System:



ANSWER: _____

12a. What is its 9th Collective Index systematic name?

ANSWER: _____

12b. What is its registry number?

ANSWER: _____

12c. What is its 8th Collective Index systematic name?

ANSWER: _____

12d. Draw an informed conclusion about the consistency and stability of substance names and registry numbers.

ANSWER: _____

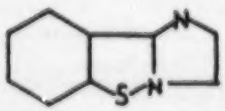
_____

Chemical Substance Names and the Index of Ring Systems

About 85% of the compounds in the Registry System contain one or more rings. Some are derivatives of dificult-to-name polycyclic, bridged or spiro ring systems. We need not despair of finding information on them if we can draw the structure. Parent compounds usually derive their names from their ring structures, so the latter make a handy way of indexing such compounds, and Chemical Abstracts Service created the Index of Ring Systems (IRS) to guide us to them by this means. The IRS is an index not to the abstracts, but rather it is a guide to the names used in the Chemical Substances Index.

IRS ----------> Chemical Substances Index

To find chemical substance names through the Index of Ring Systems, we apply the following analysis to structures:



1. Count the number of rings in the system --
   e.g., a 3-ring system
2. List the rings in order of increasing size --
   e.g., 5,5,6

3. Write the formulas of the rings in Hill System order (rings with hetero atoms precede all-carbon rings of the same size) --

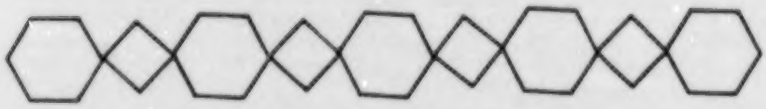
e.g., C_3NS-C_3N_2-C_6

We can then search this ring system in the Index of Ring Systems for each period. There we will find a list of compounds that satisfy the ring analysis. We may be able to recognize the correct name of our compound from the list, but probably we will need to look at the structure diagrams that appear with the names in the Chemical Substances Index.

The name of the ring system in the Chemical Substances Index provides a point of indexing entry both for the parent compound and its derivatives. The IRS for the 8th Collective Index Period is part of the Index Guide. Later editions of the IRS are in the Formula Index.

The Ring Systems Handbook is a separate publication of Chemical Abstracts Service, which contains a comprehensive list of parent ring systems. It is searched by using the same method.

13. What is the 9th Collective Index name for the parent compound of this ring system?



ANSWER: _____

13a. Give the volume and abstract number for a paper on _____ _____ derivative of the parent.

ANSWER: _____

STUDENT EVALUATION

1. What was the level of difficulty of these exercises (circle one):

   a. too easy
   b. about right for me
   c. too difficult

2. Approximately how many hours did you spend on the workbook?:

3. How useful was the experience of the workbook for your understanding of Chemical Astracts? (Please comment):

4. What parts of the workbook were particularly helpful? (Please comment):

5. What parts were least helpful?: (Please comment):

6. Please make any suggestions for improvement in the workbook that you deem useful: